# Genética e historia de las poblaciones del norte de Africa y la península Ibérica

## El análisis genético ha revelado que los amplios intercambios culturales producidos entre el Magreb y la península Ibérica no conllevaron grandes intercambios de poblaciones

E. Bosch, F. Calafell, S. Plaza, A. Pérez-Lezaun, D. Comas, J. Bertranpetit

Las poblaciones humanas se componen de individuos genéticamente distintos entre sí. Del estudio de la variabilidad de nuestra especie se ocupa la genética de poblaciones aplicada a escala mundial. A dicha disciplina le corresponde exponer la magnitud y distribución de la variabilidad genética humana.

Dos personas cualesquiera, tomadas al azar, se distinguen, en promedio, en un 0,1 % de las bases nucleotídicas que conforman su ADN. Expresado de otro modo, discrepan en seis millones de pares de bases. (El ADN humano consta de unos 3000 millones de pares de bases en cada una de las dos dotaciones haploides, una procedente del padre y otra de la madre.)

Si del individuo pasamos a las poblaciones, las diferencias observadas explican, a lo sumo, un 15 % de la disparidad genética total; un 10 % se debe a las diferencias entre grandes grupos continentales y el 5 % restante a las diferencias entre poblaciones de un mismo continente. Aun siendo pequeñas, estas últimas diferencias tienen que ver con la historia de cada población. Podemos apoyarnos en la diversidad genética entre poblaciones para reconstruir la historia demográfica.

El acervo genético de las poblaciones actuales es el resultado de la interacción entre diversas fuerzas evolutivas. Dependen éstas, a su vez, de la historia demográfica de las poblaciones, de las características intrínsecas de las regiones genómicas estudiadas y de la interacción entre genoma y factores ambientales.

Las características intrínsecas de cada región del genoma remiten a sus tasas y patrones de mutación y recombinación, así como a su modo de herencia. Se trata de parámetros que la ciencia conoce con razonable precisión.

La interacción entre la variabilidad de cada gen y el ambiente (tomado en sentido amplio, incluida, pues, la interacción con otros genes) puede promover la selección natural. Es decir, unas variantes pueden mostrarse más eficientes y verse privilegiadas por la selección, en tanto que otras pueden ser desfavorables. Lo observamos en la hemoglobina. Algunas variantes de esta proteína confieren resistencia a la malaria; la selección prima su presencia en zonas palúdicas. En consonancia con ello, el estudio de la variabilidad de la hemoglobina nos informará sobre la distribución de la malaria con mayor rigor que la historia de las poblaciones.

Conviene saber que sólo un 1,5 % de la secuencia de ADN humano llega a expresarse, es decir, determina proteínas que se sintetizan y son objeto de selección natural. Por lo tanto, la mejor estrategia para conocer la historia de las poblaciones será la que se centre en la variabilidad presente en el 98,5 % restante, cuyas probabilidades de verse afectada por la selección son mucho menores.

A partir de esta premisa podemos analizar dicha variación, explicar las diferencias genéticas neutras (polimorfismos) que encontramos entre los individuos de una población e interpretarlas en términos de historia de las poblaciones.

Disponemos de un amplio bagaje teórico, desarrollado desde los años cuarenta, gracias al cual, dada una historia demográfica, podemos predecir sus efectos sobre la diversidad genética. Podemos reconstruir

**Los autores**

E. BOSCH, F. CALAFELL, S. PLAZA, A. PEREZ-LEZAUN, D. COMAS y J. BERTRANPETIT han desarrollado la investigación que recoge su artículo en la Unidad de Biología Evolutiva de la Universidad Pompeu Fabra de Barcelona, cuya actividad se centra en el estudio de la diversidad genética en poblaciones humanas y en primates.




**1. EMBARQUE DE LOS MORISCOS** en el puerto de Vinaroz.




la historia demográfica a partir de la diversidad genética investigada en diversas regiones del genoma, que difieren en su velocidad de cambio y que permiten reconocer huellas genéticas a distintas profundidades de un tiempo pasado.

## Deriva genética

¿Cómo influye la historia en la diversidad de las poblaciones? A través de dos mecanismos básicos: la deriva genética y el flujo génico. En la deriva se engloban todos los fenómenos de cambio genético aleatorio que se dan cuando una generación transmite sus genes a la siguiente. Así como hay apellidos que prosperan y otros que se pierden en razón del número de hijos varones procreados en cada generación, las variantes genéticas (o *alelos*) pueden también cambiar de frecuencia; en ambos casos se trata de fenómenos aleatorios. Tales oscilaciones serán tanto más intensas cuanto menor sea la población, por un simple efecto de muestreo. Las desviaciones respecto a la probabilidad teórica son mayores si realizamos un número pequeño de ensayos, de la misma forma que al tirar una moneda al aire repetidas veces sólo se alcanza con seguridad la frecuencia esperada de 1/2 si se lanza muchísimas veces.

Se presenta un caso extremo de grandes cambios genéticos aleatorios, conocido por *efecto fundador*, cuando un grupo reducido de individuos establece una nueva población y se lleva consigo una muestra no necesariamente representativa de los genes de la población de origen. En la colonización sucesiva de las islas de la Polinesia, por ejemplo, se dio una secuencia clara de efectos fundadores; en el curso de la misma, un grupo limitado de individuos partía de una isla y se asentaba en la siguiente.

Las oscilaciones aleatorias de las frecuencias alélicas pueden llegar a la extinción de algunas de estas variantes. Puesto que dichas oscilaciones son más intensas en poblaciones pequeñas, se pierde variabilidad más fácilmente en éstas. A no ser que se dé una tasa de mutación extraordinaria, resulta muy poco probable que en las poblaciones pequeñas se regenere la variación perdida. Por lo tanto, al detectar una menor variabilidad genética en una población actual, podemos reconocer episodios de reducción de la población en el pasado (los llamados cuellos de botella), aunque la población actual se haya recuperado.

Además, habida cuenta de la naturaleza aleatoria de la deriva genética, las poblaciones pequeñas contiguas tenderán a diferir más entre sí que las mayores. Esta misma naturaleza aleatoria puede manifestarse de manera ligeramente diversa en regiones genómicas distintas. En consecuencia, los análisis basados en una sola región genómica pueden resultar poco fiables. Conviene siempre considerar la información procedente de un número razonable de regiones genómicas y extraer las tendencias medias.

## Flujo génico

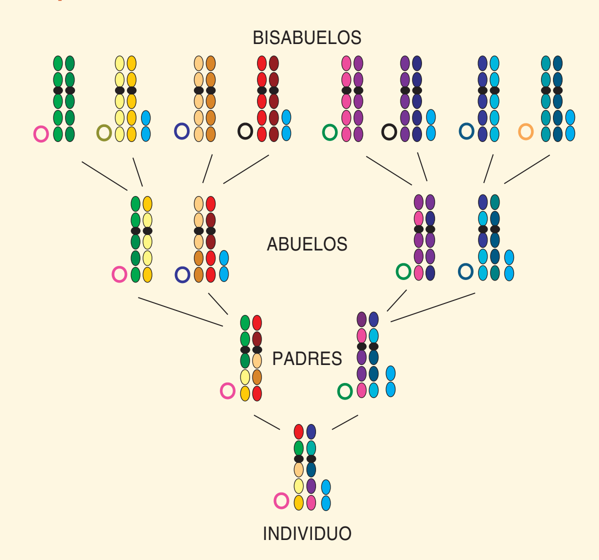
Dos poblaciones que se hayan diferenciado, por deriva, en su composición genética y entren en contacto, pueden mezclarse y dar

# Patrones de herencia: autosomas, cromosoma Y, ADN mitocondrial

Nuestro ADN se dispone en 23 pares de cromosomas. Cada miembro de un par es casi idéntico al otro en longitud y en la información que contiene; se trata de dos rasgos distintivos de cada par. Cada miembro de cada par de cromosomas nos viene de un progenitor; a cada uno de nuestros hijos le legaremos un solo miembro de cada pareja. Pero no es una transmisión fidedigna; en virtud del proceso de recombinación del material genético no heredamos el cromosoma original, sino una mezcla que contiene partes de cada miembro del par, tomadas al azar. Para ilustrarlo, la figura muestra un par de cromosomas que contiene fragmentos de distinta longitud de los cromosomas de los bisabuelos. Por eso resulta imposible predecir *a priori* de qué antepasado proviene un determinado fragmento de ADN autosómico.

Hay en el genoma dos regiones que presentan un patrón de herencia distinto. Nos referimos a los cromosomas sexuales y el ADN mitocondrial. A diferencia de los autosomas, los cromosomas sexuales (X e Y) son muy diferentes entre sí. El cromosoma Y determina la masculinidad a través de la acción de un único gen, SRY (*sex-determining region*); los cigotos con un cromosoma X y un cromosoma Y generan embriones masculinos, en tanto que los portadores de dos cromosomas X generan embriones femeninos. Por lo tanto, los varones heredan el cromosoma Y de su padre, que a su vez lo recibió del abuelo paterno, de la misma forma que se hereda el primer apellido. Lo vemos reflejado en la figura: de los cuatro bisabuelos varones, sólo el abuelo paterno del padre lega su cromosoma Y (*azul liso*) a su bisnieto.

Además de los cromosomas, que residen en el núcleo de las células, otros orgánulos contienen ADN. Se trata de las mitocondrias, que alojan decenas de copias de una pequeña molécula circular de ADN. Este ADN mitocondrial (ADNmt) se hereda por vía materna: el ADNmt del embrión procede sólo del óvulo, porque el ADNmt del espermatozoide no llega a penetrar en el óvulo. Así, en la genealogía del ejemplo, el ADN mitocondrial del individuo (*magenta*) proviene de su madre, de su abuela materna, de la madre de ésta, y así sucesivamente.




Para entender esta figura sobre cromosomas y herencia, adviértase que las barras grandes representan autosomas (cromosomas no ligados al sexo), las pequeñas representan el cromosoma Y (cuya presencia denota un varón) y los círculos, el ADN mitocondrial. De abajo arriba se esquematiza un individuo, su madre y su padre, sus cuatro abuelos y sus ocho bisabuelos.

E. BOSCH, F. CALAFELL, S. PLAZA, A. PEREZ-LEZAUN, D. COMAS Y J. BERTRANPETIT. (Adaptada de Jobling and Tyler-Smith en *Trends in Genetics*, n.º 11, págs. 449-456; 1995.)

lugar a una población con características genéticas de las dos de partida. Este fenómeno de *flujo génico*, así se le llama, se debe a la migración. La propia migración en distancias cortas, habitual a través del matrimonio, puede promover, a largo plazo, el intercambio de genes a grandes distancias. Lo observamos, por ejemplo, en las poblaciones de Asia Central, que poseen características genéticas intermedias entre las de Europa y las de Asia Oriental; su peculiar constitución genética podría explicarse por su posición central y milenios de migraciones individuales de corto alcance.

La diferenciación entre poblaciones resultante de la deriva genética se acentúa con el paso del tiempo. Para medirla disponemos de un parámetro, la *distancia genética*, que indica el grado de diferenciación entre pares de poblaciones para múltiples regiones del genoma. Si se trata de un conjunto de poblaciones, podemos representar su matriz de distancias genéticas mediante algoritmos; ofrecen éstos un paisaje genético que refleja las afinidades y diferencias dentro del conjunto poblacional. El paisaje compendia la historia de las poblaciones en términos de deriva genética y flujo génico.

Para trazar y cuantificar con razonable precisión los flujos génicos, disponemos de una nueva herramienta de análisis. Se trata de la filogeografía. Estudia ésta la genealogía del gen que ha dado origen a la variación existente dentro de una región del genoma y la distribución geográfica de dicha variabilidad.

Ante una diversidad genética dada, pensemos en una secuencia de ADN o en un conjunto de polimorfismos, la herramienta mencionada se propone reconstruir el proceso evolutivo o filogenético que ha desembocado en la diferenciación observada a partir de un antepasado común. Al plasmar conjuntamente la diversificación del gen y la de las poblaciones, podemos anclar ciertas variantes genéticas (secuencias o haplotipos) en una rama del árbol evolutivo y en un origen geográfico. Aplicando ese método se ha cuantificado la aportación por vía paterna y materna de africanos, europeos y amerindios al acervo genético de la población brasileña contemporánea.

Para datar puntos concretos de la evolución humana se puede recurrir a los microsatélites, segmentos de ADN que contienen repeticiones de breves secuencias de dos a seis nucleótidos. Son marcadores de evo-

lución rápida. A partir de un determinado acontecimiento fundador, la cantidad de variación acumulada y medible es una función de la tasa de mutación (que podemos estimar) y del tiempo transcurrido, que es la incógnita que despejaremos. Por ejemplo, se observó que la mayoría de los judíos apellidados Cohen (“sacerdote”) poseían cierto tipo de cromosoma Y. Concurre, además, que la condición de sacerdote, el apellido y el cromosoma Y se transmiten de padres a hijos exclusivamente por la línea masculina. Pues bien, de la variación acumulada en los microsatélites de este tipo de cromosoma Y se infiere un efecto fundador que operó hace unos 3000 años, coincidente con el establecimiento de una casta sacerdotal hebrea.

## El Magreb y la península Ibérica

El análisis de la diversidad genética humana en poblaciones actuales ha arrojado luz sobre numerosas cuestiones históricas, en distintas escalas temporales y espaciales. Sabemos ya que la distribución y la antigüedad de la diversidad genética a escala mundial son compatibles con un origen reciente y africano de la humanidad actual. El punto de arranque, situado en Africa, se remontaría, a lo sumo, unos 150.000 años atrás. Por tanto, ni los habitantes del yacimiento de Atapuerca ni los neandertales serían antepasados nuestros.

A escala continental, se debate la proporción de genes de origen paleolítico (hace unos 30.000 años) y neolítico (hace 10.000) presentes en los europeos actuales. Los genes, por otro lado, apuntan a una fecha antigua (unos 30.000 años) y a un origen claramente norteasiático para la colonización de América.

A escala regional, podemos abordar también algunas cuestiones abiertas sobre la historia de las poblaciones. Se cuenta aquí con la colaboración de otras disciplinas; la arqueología, la paleoantropología o la lingüística suministran a menudo hipótesis que, en la medida que impliquen distintas historias demográficas, pueden verificarse mediante el estudio de la diversidad genética de las poblaciones actuales.

Desde la genética de poblaciones podemos abordar cuestiones que atañen a la península Ibérica (España y Portugal) y al noroeste de Africa (el Magreb: Marruecos, el Sahara Occidental, Mauritania, Argelia y Túnez). ¿Se puede hablar de un origen común para ambas poblaciones a sendas orillas del Mediterráneo? ¿Quedan en las poblaciones actuales rastros de un substrato paleolítico que represente el poblamiento inicial de los antepasados de las poblaciones actuales? ¿Es el mismo substrato para ambas regiones? ¿Cuál fue la aportación de la oleada de avance neolítica? ¿Qué fracción del acervo genético magrebí proviene de la invasión árabe? ¿Podemos identificar la contribución magrebí a las poblaciones peninsulares? ¿Es el Sahara una barrera impenetrable al intercambio de genes entre poblaciones?

Para resolver esa gavilla de cuestiones sobre el poblamiento y las relaciones genéticas entre la península Ibérica y el Magreb, hemos recurrido al análisis de marcadores clásicos, microsatélites autosómicos, inserciones Alu, secuencias de ADN mitocondrial, polimorfismos de un solo nucleótido (SNP) del cromosoma Y y microsatélites del cromosoma Y en muestras de poblaciones ibéricas, beréberes del norte, centro y sur de Marruecos y del centro de Argelia, árabes marroquíes, argelinos y tunecinos, y saharauis. No se estudiaron todas las poblaciones para todos los marcadores, aunque sí se investigó extensamente un núcleo fundamental.

En el caso de los marcadores clásicos, recopilamos la información publicada por otros equipos de trabajo. En otros casos, contrastamos nuestros resultados con los obtenidos por otros autores. De la investigación realizada se desprende una descripción, que creemos ajustada, de la historia de las poblaciones norteafricanas e ibéricas.

## Orígenes remotos de norteafricanos y habitantes de la península Ibérica

Con pocas excepciones, todos los marcadores genéticos analizados muestran una separación clara entre las poblaciones magrebíes y la de España y Portugal, incluidas, sin embargo, en el rango de la variación de las poblaciones *caucasoides* (las de origen europeo, más las norteafricanas y medioorientales). Ahora bien, las distancias genéticas entre ibéricos y el resto de europeos son menores que entre ibéricos y magrebíes.

Este patrón mayoritario no se refleja en todos y cada uno de los genes analizados. Por ejemplo, en la región que lleva la información para la síntesis de los antígenos de los leucocitos humanos (HLA), los cuales definen la compatibilidad en trasplantes de órganos. Basándose en una sola región del genoma (y sometida a selección), algunos autores postularon un origen común de peninsulares y magrebíes. Pero la naturaleza aleatoria de la deriva genética y la acción de la selección pueden producir este tipo de desviaciones; para evitarlas, la interpretación debe apoyarse en la información conjunta del máximo número posible de genes y no en una sola región del genoma.

De acuerdo con nuestro análisis de marcadores clásicos, las distancias genéticas entre ibéricos y poblaciones del Oriente Medio son menores que la que existe entre ibéricos y magrebíes. Se da, además, una discontinuidad abrupta entre las orillas septentrional y meridional en el paisaje genético de la cuenca mediterránea, con una máxima pendiente en el estrecho de Gibraltar.

Estas y otras consideraciones nos llevaron a postular que en el Magreb pudo conservarse un substrato paleolítico, distinto del substrato paleolítico europeo. Supondría ello que, a diferencia de lo que parece haber sucedido en Europa, la transición al Neolítico norteafricano se produjo sin un recambio sustancial de genes. En esa hipótesis abunda la cultura capsiense del Mesolítico norteafricano, que se prolonga hacia el Neolítico adoptando las nuevas formas de producción, aunque sin la drástica ruptura que se observa en gran parte de Europa. Planteamiento que viene avalado por investigaciones con secuencias Alu y microsatélites.

Hallamos una confirmación directa de la hipótesis anterior al analizar la filogeografía del cromosoma Y.

Los linajes del cromosoma Y, definidos a partir de SNPs, presentan una genealogía muy clara y una distribución geográfica que tiende a ser restringida. Las frecuencias del mismo en una región difieren de las observadas en la otra: en la península Ibérica predominan, lo mismo que en el resto de Europa Occidental, el linaje R1b* y sus inmediatos derivados (el conjunto R1b3, R1b6 y R1b8), en tanto que dos tercios de los cromosomas Y magrebíes pertenecen al linaje E3b2*.

Tras la investigación realizada sobre la variación en el cromosoma Y ha quedado patente que el grupo de linajes R1b se encuentra sólo en Oriente Medio y Europa. Ciertos linajes (R1b3, R1b6 y R1b8) se hallan circunscritos a la península Ibérica; habrían surgido aquí a partir del haplotipo fundador, sin apenas dispersarse allende sus fronteras. El grupo de linajes R1b, por su antigüedad y distribución geográfica, se habría originado en Oriente Medio y se habría difundido por Europa con las colonizaciones iniciales del Paleolítico superior.

El linaje magrebí E3b2* se ha hallado en otras poblaciones, aunque con frecuencias mucho menores. Su antepasado más inmediato aparece entre los etíopes. Si atendemos a la acumulación de variabilidad en microsatélites asociada a E3b2*, se trataría de un linaje de más de 19.000 años de existencia. Podemos, pues, postular un escenario en que una expansión paleolítica desde Africa nororiental llevara al Magreb los antepasados del linaje E3b2*, que surgiría *in situ* después de la expansión. El linaje masculino E3b2* tiene un correlato en el linaje matrilineal mitocondrial U6, también de antepasados etíopes y difusión limitada al noroeste de Africa, aunque no alcanza las elevadas frecuencias de E3b2*. Por consiguiente, dos regiones genómicas independientes con filogeografías bien establecidas confirman la singularidad magrebí y sitúan sus raíces en el Paleolítico.

## El Neolítico: un avance paralelo

Se admite que la agricultura y ganadería comenzaron en Oriente Medio hace unos 10.500 años. Esa fase prehistórica conllevó el crecimiento y la expansión consiguiente de la población en varias direcciones. Pero se debate si dicha expansión supuso un recambio genético en las poblaciones europeas o si se conserva en la actualidad un importante substrato genético paleolítico (anterior a la expansión del Neolítico), así como el grado en que ocurrió una cosa u otra.

Si atendemos al paisaje genético europeo, advertiremos un gradiente o clina desde el sudeste hacia el noroeste. Asimismo, aparece otra clina este-oeste en el norte de Africa, desde Egipto hasta Marruecos. No es fácil dar con una explicación de dichas clinas, pues se produjeron varios movimientos migratorios en esas mismas direcciones; por ejemplo, la primera colonización del Paleolítico y el avance del Neolítico en Europa, o el Neolítico y las in-

## Tipos de polimorfismos genéticos

La variabilidad genética se presenta en múltiples formas. Con las técnicas disponibles nos es dado conocer distintas facetas de dicha variación. Podemos analizar directamente la variabilidad del ADN. Asimismo, los *marcadores genéticos clásicos* son los sistemas polimórficos detectados en los productos de expresión génica. A través de esta segunda vía, históricamente anterior a la primera, se nos revelaron los primeros sistemas polimórficos genéticos, por ejemplo, el de los grupos sanguíneos, descubiertos en 1900. Otros marcadores clásicos consisten en la variación de movilidad de proteínas plasmáticas, enzimas eritrocitarias y el sistema de antígenos leucocitarios humanos (antígenos HLA), que determinan la compatibilidad de los trasplantes.

La variabilidad o polimorfismo del ADN puede tomar varias formas. En la más sencilla, una determinada posición de nuestro genoma puede presentar dos variantes químicas (raras veces más de dos) de las cuatro posibles en las que se escribe el alfabeto del ADN (G, A, T y C). Se trata de un *SNP* (*Single Nucleotide Polymorphism*), de los que se calcula que hay unos tres millones en el genoma humano; se caracteriza por que en una posición concreta del genoma puede haber una u otra base (A o G, por ejemplo).

En vez de limitarnos a una sola posición del genoma, podemos definir un cierto segmento y determinar su secuencia en un conjunto de individuos. Esta estrategia permite, además de reconocer las variantes existentes en ciertas posiciones, descubrir también toda la variación que hay en el ADN y reconstruir su genealogía, es decir, el proceso evolutivo que dio lugar a la variabilidad observada.

Encontramos otro tipo de variación genética en los *microsatélites* o *STR* (*Short Tandem Repeat Polymorphisms*). Estas secuencias de ADN consisten en la repetición en tándem de una unidad básica de entre 2 y 6 nucleótidos de longitud. Presentan una particularidad distintiva: el número de repeticiones de la unidad básica es muy variable entre individuos. Por ese motivo se han convertido en importantes marcadores en el campo de la genética forense (identificación de personas y de la paternidad).

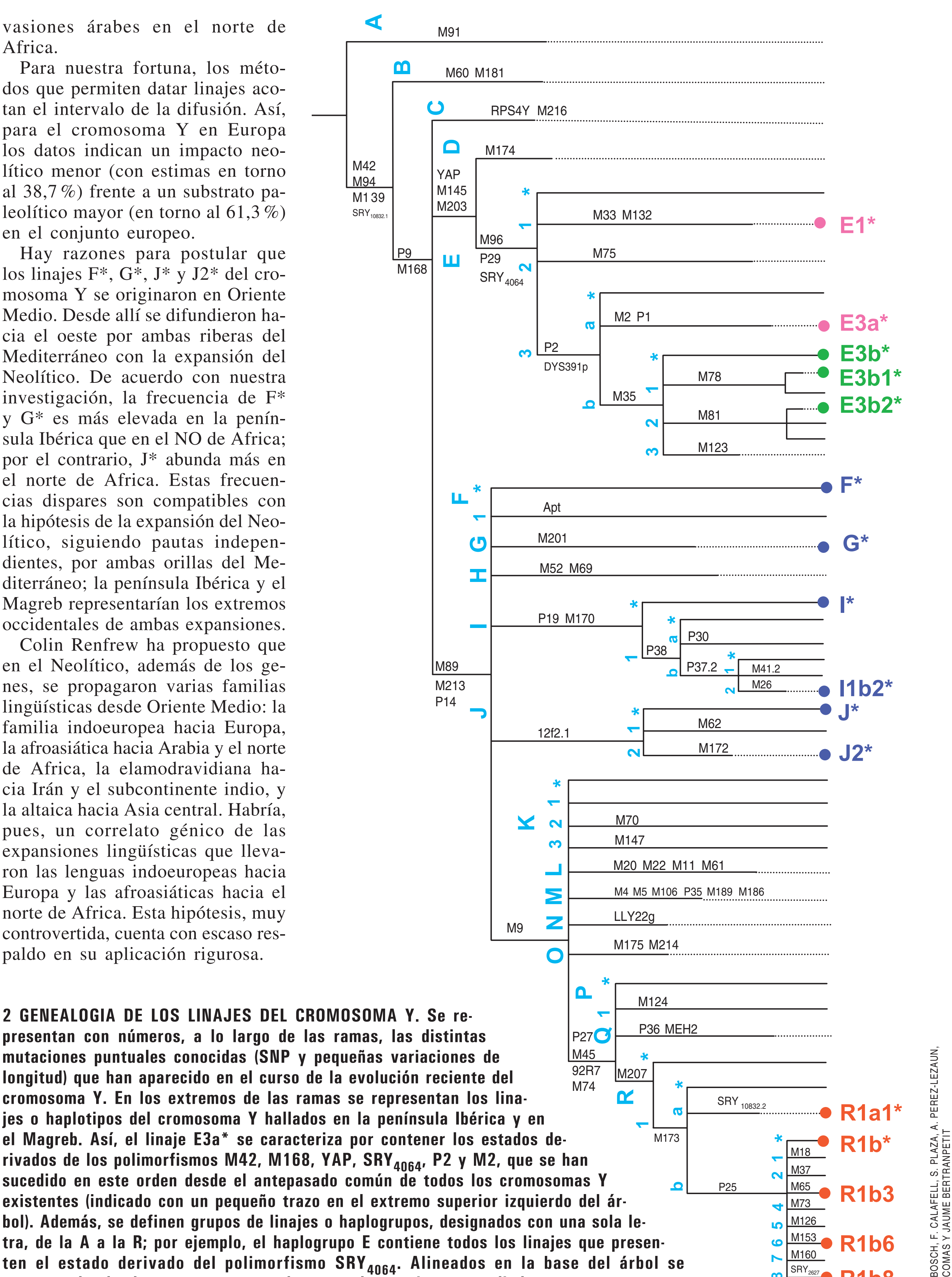
Las *inserciones Alu* constituyen, por último, otro tipo interesante de polimorfismo. Estas secuencias, de unos 300 nucleótidos de longitud, utilizan la maquinaria celular de replicación del ADN para copiarse a sí mismas y reinsertarse en otras partes del genoma. Las nuevas inserciones, lo mismo que cualquier otro polimorfismo, pueden perderse o fijarse, pero algunas de ellas se hallan en una situación intermedia. Aunque esta tercera situación se da con cierta frecuencia, es un fenómeno variable de una población a otra. Además, caso único entre los polimorfismos, conocemos bien cuál es la dirección de la evolución: normalmente, de la ausencia (estado ancestral) a la presencia (estado derivado) del elemento Alu.

vasiones árabes en el norte de Africa.

Para nuestra fortuna, los métodos que permiten datar linajes acotan el intervalo de la difusión. Así, para el cromosoma Y en Europa los datos indican un impacto neolítico menor (con estimas en torno al 38,7 %) frente a un substrato paleolítico mayor (en torno al 61,3 %) en el conjunto europeo.

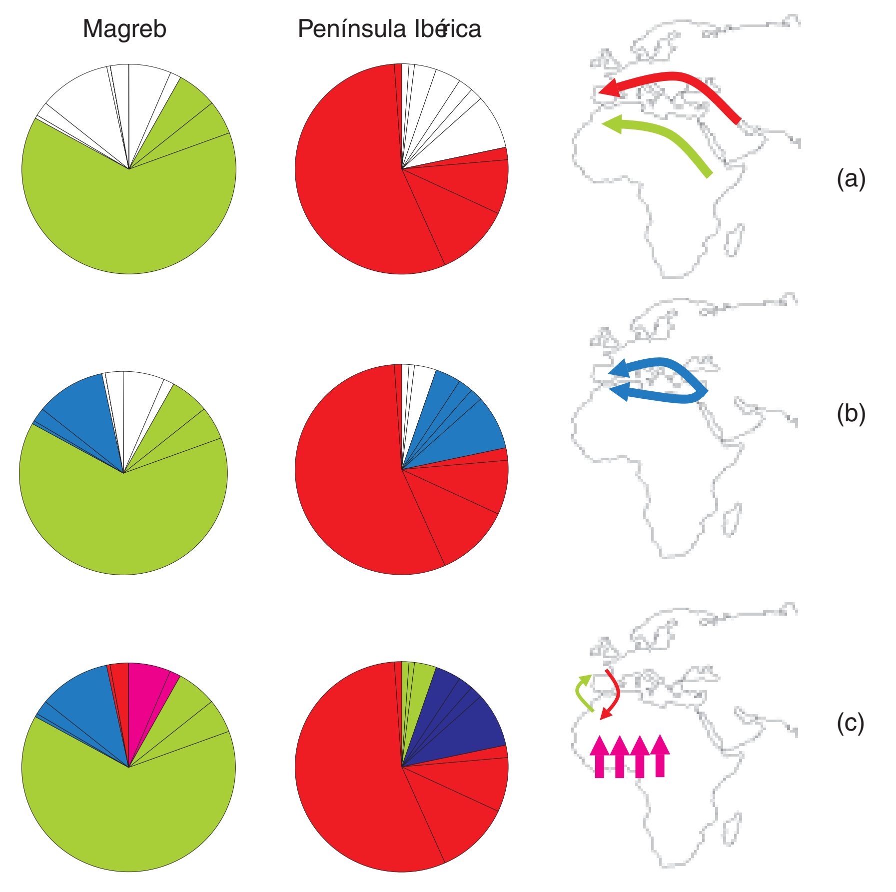
Hay razones para postular que los linajes F*, G*, J* y J2* del cromosoma Y se originaron en Oriente Medio. Desde allí se difundieron hacia el oeste por ambas riberas del Mediterráneo con la expansión del Neolítico. De acuerdo con nuestra investigación, la frecuencia de F* y G* es más elevada en la península Ibérica que en el NO de Africa; por el contrario, J* abunda más en el norte de Africa. Estas frecuencias dispares son compatibles con la hipótesis de la expansión del Neolítico, siguiendo pautas independientes, por ambas orillas del Mediterráneo; la península Ibérica y el Magreb representarían los extremos occidentales de ambas expansiones.

Colin Renfrew ha propuesto que en el Neolítico, además de los genes, se propagaron varias familias lingüísticas desde Oriente Medio: la familia indoeuropea hacia Europa, la afroasiática hacia Arabia y el norte de Africa, la elamodravidiana hacia Irán y el subcontinente indio, y la altaica hacia Asia central. Habría, pues, un correlato génico de las expansiones lingüísticas que llevaron las lenguas indoeuropeas hacia Europa y las afroasiáticas hacia el norte de Africa. Esta hipótesis, muy controvertida, cuenta con escaso respaldo en su aplicación rigurosa.




**2 GENEALOGIA DE LOS LINAJES DEL CROMOSOMA Y. Se representan con números, a lo largo de las ramas, las distintas mutaciones puntuales conocidas (SNP y pequeñas variaciones de longitud) que han aparecido en el curso de la evolución reciente del cromosoma Y. En los extremos de las ramas se representan los linajes o haplotipos del cromosoma Y hallados en la península Ibérica y en el Magreb. Así, el linaje E3a* se caracteriza por contener los estados derivados de los polimorfismos M42, M168, YAP, $SRY_{4064}$, P2 y M2, que se han sucedido en este orden desde el antepasado común de todos los cromosomas Y existentes (indicado con un pequeño trazo en el extremo superior izquierdo del árbol). Además, se definen grupos de linajes o haplogrupos, designados con una sola letra, de la A a la R; por ejemplo, el haplogrupo E contiene todos los linajes que presenten el estado derivado del polimorfismo $SRY_{4064}$. Alineados en la base del árbol se muestran los haplogrupos representados en ambas regiones estudiadas.**

E. BOSCH, F. CALAFELL, S. PLAZA, A. PEREZ-LEZAUN, D. COMAS Y JAUME BERTRANPETIT




**3. HISTORIA DE LA POBLACION y linajes del cromosoma Y en el Magreb y la península Ibérica. (*a*) La primera colonización del Paleolítico se da independientemente en ambas regiones; introduce en el Magreb (*verde*) los linajes E3b*, E3b1* y E3b2*; en la península Ibérica (*rojo*), los linajes R1a1*, R1b* (que ulteriormente dio lugar a R1b8), R1b3 y R1b6. En los diagramas de sectores se muestra la frecuencia de dichos linajes en cada población. (*b*) La expansión del Neolítico, desde el Creciente Fértil y en paralelo por ambas riberas del Mediterráneo, aporta los linajes F*, J*, J2*, I* e I1b2* (*azul*). (*c*) Los fenómenos migratorios implican flujo génico desde la península hacia el Magreb (*en rojo*), en sentido contrario (*verde*), y desde más allá del Sahara hacia el Magreb (*malva*).**

## Arabes y beréberes

Los beréberes (o *imazighen*) constituyen los descendientes directos de una población ancestral que se extendía por gran parte del norte de Africa, desde Egipto hasta Senegal. Suman hoy 20 millones de personas, dispersas en pequeñas minorías de Egipto, Libia y Senegal. Sin embargo, muchos habitantes de Túnez, Argelia y Marruecos se definen a sí mismos como tales y hablan alguna de la veintena de lenguas beréberes, una rama de la familia afroasiática. El resto de la población habla y se considera árabe, y se supone descendiente de las invasiones, que desde el siglo VII y con especial intensidad en el XI, llevaron el Islam desde la península Arábiga hasta el Magreb.

Así las cosas, podemos plantearnos si las invasiones árabes implicaron una aportación demográfica significativa o si, por el contrario, una elite numéricamente limitada pero culturalmente prestigiosa consiguió difundir una nueva lengua y religión, sin que ello conllevara una aportación de genes notable. Para resolver tal disyuntiva hemos de acudir al análisis genético de las poblaciones árabes y beréberes.

El análisis de gran cantidad de marcadores (inserciones Alu, microsatélites autosómicos y polimorfismos del cromosoma Y) nos revela una llamativa ausencia de diferencias entre poblaciones árabes y beréberes. Sólo el ADN mitocondrial separa de los beréberes a los árabes argelinos y tunecinos (pero no marroquíes). Debemos concluir, pues, que la arabización del Magreb fue un fenómeno básicamente cultural, en que una reducida elite impuso su lengua y religión, sin que hubiera cambios sustanciales en la población local, incluso la actualmente arabófona.

## Más allá del Sahara

La comparación de la diversidad genética con la hallada al sur del Sahara permite rastrear el flujo génico transahariano. En el estudio de las inserciones Alu se advierte con nitidez que las poblaciones más al sur de nuestra zona de trabajo (saharauis y beréberes del sur de Marruecos) muestran distancias genéticas más cortas con las poblaciones subsaharianas que las que se dan entre subsaharianos y poblaciones del norte del Magreb. Tal comprobación nos induce a pensar en un gradiente de flujo génico subsahariano; en el curso del mismo, las poblaciones del sur del Magreb habrían recibido una mayor aportación de genes subsaharianos, lo que, dada su posición geográfica y el conocido comercio de esclavos, parece verosímil.

Las regiones genómicas con una filogeografía bien establecida permiten cuantificar la aportación subsahariana. Así, en el norte de Africa aparecen en bajas frecuencias (un 8 % en conjunto) los linajes E1* y E3a* del cromosoma Y, de origen subsahariano; no se han hallado en la península Ibérica.

En el caso del ADN mitocondrial, son de origen subsahariano los linajes L1, L2 y L3, que constituyen una media del 25 % de los linajes magrebíes (con un rango entre 3 % en rifeños y 40 % en mauritanos). En la península Ibérica, presentan una frecuencia media del 3 %, oscilando entre su ausencia en vascos y un 6 % en portugueses del centro. En el caso de la Península, es difícil decidir si estos linajes proceden directamente de allende el Sahara, traídos con la trata de esclavos, o si, dada su frecuencia en el Magreb, llegaron a Iberia vía con-

E. BOSCH, F. CALAFELL, S. PLAZA, A. PEREZ-LEZAUN, D. COMAS Y JAUME BERTRANPETIT

tactos a través del estrecho de Gibraltar.

De la comparación entre los datos del ADN mitocondrial y los del cromosoma Y se desprende que los linajes subsaharianos heredados por vía materna se hallan a una frecuencia más elevada en magrebíes e ibéricos que los linajes paternos, lo que indicaría una diferencia entre sexos en la movilidad de los individuos desde el sur del desierto del Sahara. Esa observación genética debe contrastarse con datos sociales de movilidad y comercio de esclavos.

## Tráfico en el estrecho de Gibraltar

¿Qué decir, por último, de las relaciones entre las poblaciones magrebíes y las peninsulares? Dejamos constancia al principio de la nítida separación entre ambas poblaciones, debido, probablemente, a un substrato paleolítico distinto. Ese hito temporal permite la detección del flujo génico a través del estrecho, así como su cuantificación a partir de linajes del cromosoma Y y del ADN mitocondrial.

Las personas y, si se reproducen, sus genes han cruzado el estrecho de Gibraltar con distinta intensidad a lo largo de la historia. En algunos períodos, dicho flujo aumentó. Además, se trata de una corriente bidireccional, pues también se dio un flujo génico de la Península al Magreb. De norte a sur, cruzaron el estrecho romanos, vándalos, judíos y moriscos. Los dos últimos grupos podrían haber difundido genes ibéricos. Del sur llegaron a Hispania los cartagineses; en el 711 arribaron los árabo-beréberes. En el siglo XII vinieron oleadas de almohades, almorávides y benimerines. Los datos genéticos no permiten precisar cuándo se produjo el tráfico. Sólo podemos descubrir el conjunto acumulado de los intercambios genéticos.

Como hemos comentado, el linaje E3b2* del cromosoma Y se originó en el Magreb, donde constituye unos dos tercios del total. En España y Portugal, su frecuencia se estima alrededor del 6 %, con mínimos en el País Vasco y Cataluña y máximos en Extremadura y Andalucía occidental. Dado que el flujo génico del Magreb hacia la Península acarrearía otros linajes, subestimaríamos la contribución genética magrebí de la Península si sólo consideráramos E3b2*; corrigiendo a tenor de la frecuencia de E3b2* respecto al total de linajes del cromosoma Y, la contribución norteafricana al acervo genético ibérico se puede estimar en un 8 %. Al estudiar la variación de microsatélites dentro de este linaje, se observa una estrecha similaridad entre los haplotipos ibéricos y los magrebíes. La variación observada en haplotipos peninsulares pudo haberse acumulado en un intervalo temporal que la hace compatible con las entradas del siglo VIII y, sobre todo, con las del siglo XII.

Por lo que respecta al ADN mitocondrial, encontramos un equivalente de E3b2* en U6, de origen magrebí, aunque menos frecuente. Se halla en un 10 % de los magrebíes y en un 1,5 % de los habitantes de nuestra península. La ausencia de linajes maternos específicamente norteafricanos a frecuencias moderadas o elevadas dificulta la estimación de la contribución femenina magrebí a la península; aunque existe, obviamente.

En un sentido inverso, los cromosomas Y del grupo R1b que hay en el Magreb pueden ser de origen europeo, si bien no podemos precisar que fuera específicamente ibérico. Su frecuencia, del 2,8 % en norteafricanos, alcanza el 78,4 % en ibéricos; ello supone una contribución europea del 3,6 % al acervo genético magrebí. Por lo que respecta al ADN mitocondrial, el linaje V, de origen europeo, se encuentra en una frecuencia del 6,8 % en norteafricanos.

## Conclusiones

La historia nos recuerda las intensas relaciones culturales y sociales que han existido, a lo largo de los siglos, entre el Magreb y la península Ibérica (España y Portugal). Elites o pueblo llano, mercaderes o guerreros, portaban una lengua, una religión, una cultura, en definitiva. Pero hasta ahora se nos mostraba esquivo el impacto demográfico ejercido por esos flujos, que los datos genéticos nos muestran existente pero moderado. Y todavía se debaten los movimientos relacionados con la expansión islámica.

Los datos genéticos, con sus limitaciones, han permitido trazar un primer marco comparativo entre ambas orillas del Mediterráneo, con la reconstrucción consiguiente de la historia e intercambios mutuos de sus poblaciones. La genética aporta el marco de la historia demográfica, en cuyo interior hemos de identificar las pruebas de intercambios suministradas por otras disciplinas.

Los cambios y sustituciones en el credo religioso, en la lengua o en los perfiles de las excavaciones arqueológicas nos hablan de interrelaciones y desplazamiento culturales. El alcance demográfico de los procesos demográficos asociados a esas transformaciones culturales hallan un correlato genético, cuya magnitud se va desentrañando merced al avance en el conocimiento del genoma. Es una de las múltiples sorpresas que la biología actual nos depara.

## Bibliografía complementaria


GENETIC STRUCTURE OF NORTHWESTERN AFRICA REVEALED BY STR ANALYSIS. E. Bosch, F. Calafell, A. Pérez-Lezaun, J. Clarimon, D. Comas, E. Mateu, R. Martínez, B. Morera, Z. Brakez, O. Akhayat, A. Sefrani, G. Hariti, A. Cambon-Thomsen y J. Bertranpetit en *European Journal of Human Genetics*, n.º 8, págs. 360-366; 2000.

ALU INSERTION POLYMORPHISMS IN NW AFRICA AND THE IBERIAN PENINSULA: EVIDENCE FOR A STRONG GENETIC BOUNDARY THROUGH THE GIBRALTAR STRAITS. D. Comas, F. Calafell, N. Benchemsi, A. Helal, G. Lefranch, M. Stoneking, M. A. Batzer, J. Bertranpetit y A. Sajantila en *Human Genetics*, n.º 107, páginas 312-319; 2000.

HIGH RESOLUTION ANALYSIS OF HUMAN Y-CHROMOSOME VARIATION SHOWS A SHARP DISCONTINUITY AND LIMITED GENE FLOW BETWEEN NORTHWESTERN AFRICA AND THE IBERIAN PENINSULA. E. Bosch, F. Calafell, D. Comas, P. J. Oefner, P. A. Underhill y J. Bertranpetit en *American Journal of Human Genetics*, n.º 68, págs. 1019-1029; 2001.